智能集约高效的融媒采编系统技术应用探索

冯恩达
（新华报业传媒集团，江苏 南京 210000）

摘　要：《关于加快推进媒体深度融合发展的意见》为融媒体向智媒体发展指明了方向，本文通过对现有融媒体采编系统的痛点研究，探索智能能力组合应用、高效内容生产助手、高效的内容生产联动等智能选型和应用方法，以适应新传播形式对内容生产体系提出的更高要求。

关键词：融媒体采编；智能助手；集约高效；对话系统　　**中图分类号：**G237.6　　**文献标识码：**A
文章编号：1671-0134（2021）10-157-04　　DOI：10.19483/j.cnki.11-4653/n.2021.10.047
本文著录格式：冯恩达 . 智能集约高效的融媒采编系统技术应用探索 [J]. 中国传媒科技，2021（10）：157-160.

导语

2020年9月中共中央办公厅、国务院办公厅印发《关于加快推进媒体深度融合发展的意见》[1] 对媒体融合的下一步发展提出了明确的总体要求，更好地为各级媒体由融媒体向智媒体发展指明了方向。近年来，5G、大数据、云计算、物联网、区块链、人工智能等技术高速的发展也使得内容生产的形式与传播模式有了新的变化，催生了一些新的媒体采编发方式。如何利用好这些技术发展成果，如何将其与现有的系统建设整合融合，革新现有的内容生产体系成为许多报业媒体接下来的新课题。

1. 融媒体系统建设现状与问题

1.1 主流媒体采编系统建设概况

各大传统媒体的采编系统大致经历了传统采编、中央厨房、融媒体采编三个发展阶段。各级媒体基本已经具备了内容编辑系统、签发管理系统、媒资库等与内容生产相关的信息系统基础，迈开了融合发展的步伐，直至 2019 年 12 月新华社试点建成首个智能化编辑部，2020 年 12 月《人民日报》发布创作大脑，两大央媒揭开了利用人工智能再造内容生产流程的序幕。

1.2 融媒体系统建设中的痛点

目前融媒体转向智媒体的项目建设一般采用整体设计分期建设的方式，受限于预算、工期、子项目划分等现实条件的约束，为了实现每个时间节点系统设计功能或多或少会出现在某些类似的智能力上重复投入，跨项目能力无法相互调用的情况。同时，采用的智能能力能否快速根据融媒体业务需求进行重构、再训练，以及优化训练所需的资源消耗也是限制目前智能能力运用灵活性的一大因素。

随着融合的不断深入，融媒体系统建设不仅仅局限于技术建设项目范畴，更是人、财、物的整合融合。利用智能驱动，构建积极、开放、高效的人才队伍培养与组织方式结合有新技术加持的内容生产方式才能够完成媒体深度融合的总目标。

2. 集约高效的智能能力应用

2.1 智能能力组合应用

语音识别能力是现在智能中台具有的一种常见智能能力，该能力通过提取语音信号中的声学特征与声学模型和语言模型进行智能匹配，将语音信号转换成文本内容。目前，国内的科大讯飞开放平台、百度 AI 开放平台、阿里云等技术公司在此方面可以提供稳定可靠的智能服务。

根据调研，在融媒系统中单纯的语音识别能力多用于采访录音的转文字、视频音轨转文字等简单的内容创作辅助工作任务，如果将此能力与媒资检索入口进行结合，就可以实现以语音输入内容为检索条件的快速检索；将此能力进行实时性拓展再与移动端的采编模块进行结合，就可以实现语音输入编写稿，提高在移动端编写稿件的效率；将拓展了实时性的能力配合时间轴属性，就得到了视频智能字幕的智媒体工具……随着智媒体时代的临近，对智能能力的复用必然越来越普遍，每个子系统单独配置易造成资源浪费，通过将智能能力以微服务形式封装到智能中台里可以为日后媒体业务转型发展打下灵活高效的基础。

2.2 高效的智能助手

“大数据 + 智能”能力的运用改变了目前采编人员的工作方式和负荷，新技术一方面给采编工作带来前所未有的便捷，同时由于素材、信息、形式、选择等大大增多，采编人员也面临新的挑战。通过充分利用人工智能技术发展成果的智能助手可以以任务为导向帮助采编人员更

高效的获取相关服务、素材及信息，提高采编任务的响应速度和完成效率，从容应对将来响应式内容生产的市场新需求。

2.2.1 智能助手的对话优化方案

智能助手通常是通过将用户输入转化为有效的服务或内容输出来帮助提高效率的。目前的智能助手较多的应用场景是网上智能客服，该类客服多采用任务型的对话系统来实现对用户需求自动回复、产品推荐、服务唤起等功能。经过研究，目前市场上使用较为普遍的是基于检索式的人机对话系统，此类系统对用户提出的问题理解往往不够，比较依赖人工事先在系统中设置的响应模板，通过与数据库中的关键词进行匹配筛选出回复语句或拉起服务。优点是可以标准化给用户以答案，缺点也很明显，几轮对话之后很可能需要选择“此答案没有帮助”或者“人工客服”的按钮。由此看来，在融媒体系统场景中此类解决方案并不是很适用，本文的论述将围绕充分利用人工智能深度学习发展成果的生成式人机对话系统进行。

2.2.2 提升智能助手响应能力的方法研究

2.2.2.1 在人机对话中加入分类意图识别，优化交互体验，提高服务抵达效率

由于新闻内容生产系统中常用的术语具有一定的专业性，在人机对话过程中对用户的输入在垂直领域内进行分类可以强化对用户意图的识别准确率。模型的训练可通过基于条件序列对抗网络的意图识别[2]方法，先将用户输入进行四步预处理：①对输入进行分词并将词向量、词性标签、词频、主题等特征进行提取；②特征值拼接并向量化；③通过 RBF 函数计算向量间距，将间距近的特征向量归为同一类；④少量的人工干预复核。在对话生成的过程中使用无监督的对抗网络结构如图 1，相较有监督的方式减少模型训练的成本。结构中的生成器可通过少量的语料信息和标注信息生成足够的训练素材，可选用对长序列处理很友好的 LSTM 来完成。[3]生成器对无标注的数据初始设定均值为 0 的标准正态分布，有标注的数据采用 one-hot 编码形式表示，经过生成器将得到生成样本的概率分布。分类器中的二分类器对真假语料进行甄别，类别分类器负责通过分布概率将样本进行分类。考虑到应用场景中人机会话的输入长度通常不会太长，采用 TextCNN 卷积神经网络通过超参数调整的方式可以快速地完成短文本意图分类的工作。生成器将学习到的语句特征大量生成训练用的输入可帮助分类器快速达到理想的效果。

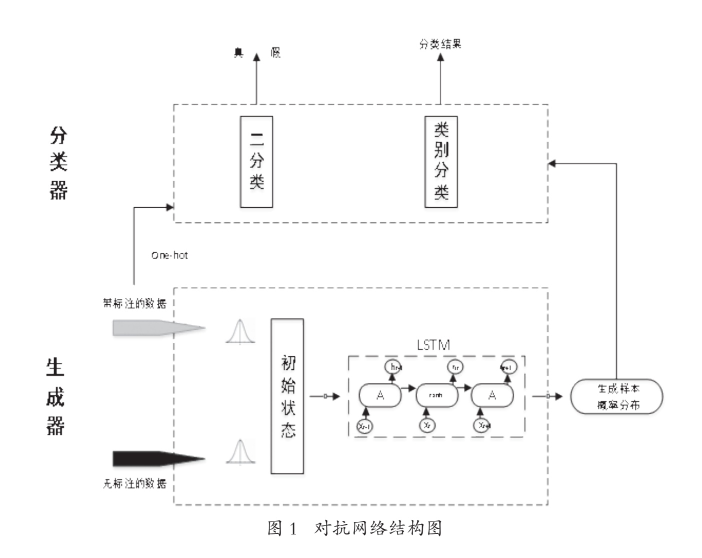


图 1 对抗网络结构图

2.2.2.2 在人机对话过程中利用媒体知识图谱支撑内容决策提高机器回复内容生成的质量

采编人员在融媒体中心这一特定场景中与机器进行对话往往是涉及系统中的资源、功能和服务部分，且报业媒体由于经过了常年累积已经具备了相当体量的行业知识积累，通过对生产系统中已有的结构化、半结构化和非结构化数据进行处理即可形成可服务于人机对话系统的知识图谱。对已有的结构化数据如人员信息（姓名、性别、年龄、职务、岗位、负责采编的方向、写过的稿件等）可利用 D2R 技术将来在中央信息厨房系统中的关系型数据转换为 RDF 数据，[4]将结构化数据的查询结果转换成三元组的形式完成知识的抽取。

由于融媒体系统中的数据来源通常不是单一源，在完成知识抽取后还需对这些异构数据在统一的架构下进行数据清洗、实体对齐、属性值决策和关系构建从而实现知识融合。在融合后的知识基础上通过实体相似度计算将同类的实体进行归类，并对归类后的知识通过 ANN 或 PRA 进行推理以构建新的实体关系，再通过质量评估筛选出可信度高的部分形成高质量的知识图谱并存于 Neo4j 数据库中。对半结构化、非结构化数据需要通过更为复杂智能模型知识抽取过程才能合并到知识图谱中。

由于此处讨论的智能助手主要是服务于内容生产的，所以人机对话的目标任务自带了范围约束，在某种程度上是对实现复杂度进行了降维。首先从完整的知识图谱中过滤出更小范围的子图集合，过滤的方法可以选用模糊检索。在缩小了范围的子图上进行基于策略梯度算法的强化学习选择出最优的知识。整体结构如图 2。

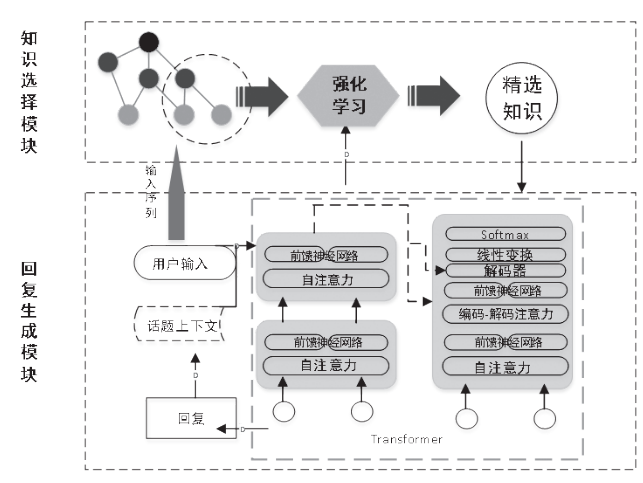


图 2　基于强化学习的网络结构图

知识选择模块通过同时将三元组知识、对话主题信息及其上下文信息同时作为输入，通过每一轮次的人机对话内容以初始话题为出发点在上述子图集中通过调整梯度优化算法的网络参数寻找最优的输入——目标拟合路径，最后将用户引导到意图识别出的目标话题所指的实体上。并通过已知的初始话题和目标话题序列以及整个对话的其他数据来进行强化学习寻找到利用知识图谱中的哪些知识来构建最优的回复话题。

前期调研中大多数的人机对话过程都被当作新的对话来进行对待，机器生成的回复没有充分利用前面对话的历史信息，实际上一个拟人的对话过程通常会结合历史对话情况做出下一次回复的。[5] 故回复生成模块采用 Transformer 网络对输入文本进行编解码的方式来训练，利用网络的 self-attention 层 [6] 将每个 encoder 在编码过程中查看输入序列中的上下文和距离相对比较远的单词。相比使用流行的 RNN 网络而言，Transformer 不需要循环且不存在时间序列上的依赖，能够并行处理序列中所有的单词和符号，它的训练速度要快的多，且生成的回复结果也更优质。在训练过程中不将知识选择和回复生成两个模块独立开训练而采用联合训练，以利用强化学习的方式帮助模型更好地利用用户输入和知识图谱的输出结果生成更优的回复。

3. 智能高效的生产联动

3.1 智能驱动的采编

目前传媒市场的供求关系正在技术的影响下发生着微妙的变化，以前用户看什么基本依赖媒体单位生产什么，现在随着万物皆媒时代的慢慢临近，用户内容创作已经逐步有了市场，自媒体行业发展如雨后春笋，与此同时问题也不断爆出，虚假新闻等缺乏新闻职业精神的作品时不时就成了网上的爆点。在此现象的背后刚好戳中了传统媒体内容生产惯性的痛点，内容编发速度不够快、内容的关注点和百姓关注点脱离、内容表现形式不能抓住百姓胃口等。《关于加快推进媒体深度融合发展的意见》中提出主力军全面挺进主战场，占领新型传播高地，作为主流媒体首先必须同步用户需求，新的传播环境下用户贪“新”图“快”的要求更甚，这不是简单地建设稿件编发平台就能解决的，还需利用智能能力让用户参与内容生产，响应式新闻产品刚好切中了用户的需求，这样的新闻可以获得用户更多共鸣。

3.1.1　用户间接参与内容生产

在稿件编辑时通过对用户已输入的内容进行语义理解，结合利用大数据汇集能力将以往和当前用户对类似主题内容的访问数据（阅读、转发、评论、点赞、停留时间等）通过智能分析后反馈给编辑人员，并可依据传播力历史信息和用户喜好数据预测内容将来传播趋势。用户在无感交互的情况下参与内容生产的引导，帮助采编人员生产出更符合受众胃口、更能引起用户兴趣的内容。

3.1.2　用户直接参与内容生产

有研究表明，网民每天阅读新闻使用最多的渠道排行是微信、抖音、今日头条，其中各渠道的特点存在社交属性和强联系性。用户开始注重新闻传播过程中参与感，目前直播的火爆正是因为用户在观看直播的过程中通过和主播互动获得了参与感，通过与主播对话甚至可以改变直播的主题内容，发表留言会得到投至屏幕上的反馈，甚至可以通过弹幕和同时在看的网友进行交流。响应式新闻就是一种以用户参与为特点的内容生产方式，如图 3。用户通过对新闻客户端的智能助手提出明确要求

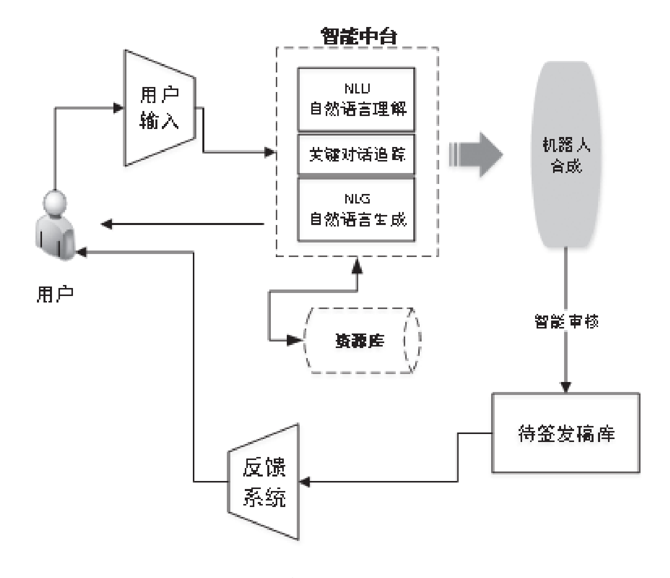


图 3　用户参与的响应式内容生产

（语音、文字、视频），客户端将信息传回内容生产平台，通过自然语义理解、视频主题捕捉等智能能力将用户需求转换成内容生产要求。联合调度智能媒资库、写稿机器人、智能剪辑等智能中台能力快速生产对应主题的文字、图片、视频等新闻内容，最后通过智能审核完成初审进入待签发稿库。

3.2 资源整合融合

新颖快速的内容生产光靠单纯的采编系统是不够的，与内容生产相关的一系列子系统都需要联动起来形成顺应新传播形态的内容生产体系才能更好应对市场的挑战。报业传媒行业依靠版面经营的黄金时代已经过去，新媒体运营时代对活动组织、策划、宣传提出了高的要求，需要联合采编业务系统、全媒体考核系统、人力资源系统等生成人员画像，针对不同的主题报道活动利用系统智能能力甄选合适的人员组成工作组，管理团队、内容团队、技术团队协同共鸣，实现全域跨媒体的人财物联动，将媒介资源和生产要素整合融合。

结语

在 5G、AI、大数据等技术快速发展的当下，融媒体向智媒体发展是必然趋势，各级媒体对智能能力的部署条件并不相同，找到适合自身定位和资源配置的智能应用方案将是每个媒体单位关注的焦点。新的传播形态将不断向报业的内容生产体系发起挑战，未来的内容生产系统会在能力模块的引入和架构调整的灵活性方面提出更多在集约化、高效化、智能化的要求。媒

参考文献

[1] http://www.gov.cn/zhengce/2020-09/26/content_5547310.htm

[2] 周新宇，基于神经网络的汉语 asr 识别对话短文本意图匹配 [D]. 北京：北京邮电大学，2020.

[3] M.Mirza, and S.Osindero ,“Conditional generative adversarial nets”, arXivreprint arXiv:1411.1784,2014.

[4] 孙国梓 . 基于知识图谱的多轮问答系统 [C] .2020.4.

[5] 徐聪 . 基于深度学习和强化学习的对话模型研究 [D]. 北京：北京科技大学，2019.

[6] Vaswani A,Shazzer N ,Parmar N,et al. Attention is all you need [C].Advances in neural information processing systems. Long Beach,2017:5998-6008.

作者简介：冯恩达（1985-），男，江苏无锡，电子信息工程中级职称，研究方向：信息化系统后端开发、智能能力应用。

（责任编辑：张晓婧）

（上接第85页）

一个信息传播的渠道，更应该是营销平台的拓展。虽然平台有不少的用户数，但由于线上产品无法抓住用户的需求痛点，线上营销薄弱，无法直达需求用户，形成供求双方信息不对称，平台的用户资源和内容资源难以实现变现。

对策：面对这样的困境，在平台的日常运营中，要注重发现用户的需求点，积极策划产品来满足用户潜在需求。在产品营销方面，要将传统出版产品与线上新产品结合起来，做好整体营销。在平台的建设中，不仅要重视平台的内容生产，还要做好多平台互动运营；不仅要重视平台用户数量的增长，还要重视平台变现能力的提升。

结语

以上是笔者在实际工作中的一些心得体会，在一定程度上为少儿报刊融媒体平台的建设提供了参考。在互联网技术迅速发展的今天，在出版行业融合发展的当下，助学类少儿报刊的产品形态和营销模式正在改变，少儿报刊与新媒体融合发展任重道远。新媒体发展日新月异，近几年，微信公众平台的发展已经进入平台期，短视频已经站在了新媒体的舞台中央。如何在数字出版和新媒体运营上取得突破，是少儿报刊需要不断思考、实践和总结的问题。媒

参考文献

[1] 孙真福 . 教育出版融合发展探究 [M]. 南京：江苏凤凰教育出版社，2020.

[2] 杨琦 . 新媒体时代如何做好少儿期刊微信公众号的运营 [J]. 领导科学论坛，2020（13）：75-77.

作者简介：李航（1985-），男，安徽安庆，江苏凤凰报刊出版传媒有限公司《时代学习报》编辑部数字发展部主任，出版中级，研究方向：少儿类报刊出版、数字出版。

（责任编辑：张晓婧）